International Journal of Computer Science Engineering
and Information Technology Research (IJCSEITR)
ISSN(P): 2249-6831; ISSN(E): 2249-7943
Vol. 4, Issue 2, Apr 2014, 115-126

VIDEO COPY DETECTION USING FINGER PRINTING WITH FAST IMAGE PROCESSING

LAXMI GUPTA[1], M. B LIMKAR[2], SANJAY M HUNDIWALE[3] & SONALI JADHAV[4]

[1]M.E Student, EXTC, ARMIET College of Engineering, Sapgaon, Mumbai, Maharashtra, India

[2]Associate Professor, Department of Electronics, Terna College of Engineering, Nerul, Mumbai, Maharashtra, India

[3]Associate Professor, Department of EXTC, ARMIET, College of Engineering, Sapgaon, Mumbai, Maharashtra, India

[4]Assistant Professor, Department of Electronics, Terna College of Engineering, Nerul, Mumbai, Maharashtra, India

ABSTRACT

Video copy detection has been actively studied in a wide range of multimedia applications. A video copy detection system is the process of detecting illegally copied videos by analyzing them and comparing them to original content. It is based on content fingerprinting and can be used for video indexing and copyright applications. This system is based on a property of fingerprint extraction algorithm followed by a fast approximate search algorithm. In this, a unique signature is created for the video on the basis of the video's content. The fingerprint extraction algorithm extracts compact content-based signatures from special images constructed from the video. Each such image represents a short segment of the video and contains temporal as well as spatial information about the video segment. These images are denoted by temporally informative representative images. To determine the query video, the fingerprints of all the videos in the database system are extracted and stored in advance. The fingerprint can be compared with other videos' fingerprints stored in a database. The search algorithm searches the stored fingerprints to find close enough matches for the fingerprints of the query video. The proposed fast approximate search algorithm facilitates the online application of the system to a large video database of tens of millions of fingerprints, so that a match is found in a few seconds. The proposed system is tested on a database of different videos in the presence of different types of distortions such as noise, changes in brightness/contrast, frame loss, shift, rotation, and time shift which emphasize the robustness and discrimination properties of the copy detection system.

KEYWORDS: Fingerprinting, Signatures, Temporal Information, Video Copy Detection, Video Indexing, Videos

INTRODUCTION

Nowadays thousands of videos are being uploaded to the internet and are shared every day. Out of these videos, considerable numbers of videos are illegal copies or some videos are manipulated versions of existing media. Therefore copyright management on the internet becomes a complicated process.

Today's widespread video copyright infringement calls for the development of fast and accurate copy-detection algorithms. To detect infringements, there are two approaches. First is based on watermarking and other is based on Content Based Copy Detection (CBCD).. Watermarking is used to detect whether images are copied or not. The first limitation of watermark is that if the original image is not watermarked, then it is not possible to know whether other images are copied or not. The second drawback of watermarking is that the degree of robustness is not adequate for some of the attacks that encounter frequently. To overcome limitations of watermarking another technique is developed called as

Content Based Copy Detection (CBCD).communication technologies, such as adoption of more efficient multimedia coding standards and the astounding increase in data transfer rates. The primary aim of Content Based Copy Detection (CBCD) is "the media itself is the watermark", that is, the media (video, audio, image) contains enough unique information that can be used for detecting copies. The key advantage of Content Based Copy Detection (CBCD) over watermarking is the fact that the signature extraction can be done after the media has been distributed. Content Based Copy Detection finds the duplicate by comparing the fingerprint of the query video with the fingerprints of the copyrighted videos.

PROBLEM STATEMENT

Research that began a decade ago in video copy detection has developed into a technology known as "video fingerprinting". The process of extracting a fingerprint from the video content is referred to as fingerprinting the video or video fingerprinting. There is an obvious analogy to human fingerprint and video fingerprinting. The analogy extends to the process of subject identification by fingerprint: first, known fingerprints must be stored in a database; then, a subject's fingerprint is queried against the database to match. Content Based video Copy Detection system can be used for video indexing and copyright applications. Previous fingerprinting extraction methods can be applied to specific videos, some can be applied only to large video sequences, and some contain only spatial information. Therefore spatio-temporal fingerprinting extraction algorithms are designed. Proposed a fingerprint extraction algorithm Temporally Informative Representative Images - Discrete Cosine Transform (TIRI-DCT) extracts compact content based signatures from special images constructed from the video. Each such image represents a short segment of the video and contains temporal as well as spatial information about the video segment.

These images are denoted by temporally informative representative images. To find whether a query video (or a part of it) is copied from a video in a video database, the fingerprints of all the videos in the database are extracted and stored in advance. The search algorithm searches the stored fingerprints to find close enough matches for the fingerprints of the query video. The proposed fast approximate search algorithm facilitates the online application of the system to a large video database of tens of millions of fingerprints, so that a match (if it exists) is found in a few seconds.

LITERATURE SURVEY

Previous video fingerprint extraction algorithms are classified into four groups as color-space-based fingerprints, temporal fingerprints, spatial fingerprints and spatio-temporal fingerprints.

Color-Space-Based Fingerprints

Color-space-based fingerprints are among the first feature extraction methods used for video fingerprinting. They are mostly derived from the histograms of the colors in specific regions in time and/or space within the video. Advantages of color histograms are efficiency and insensitivity to small changes in camera viewpoint. Color histograms are frequently used to compare images Color histograms are computationally trivial to compute.

Color histograms also have some limitations. A color histogram provides no spatial information; it merely describes which colors are present in the image, and in what quantities. In addition, color histograms are sensitive to both compression artifacts and camera auto-gain.

The first disadvantage of color-space-based fingerprint is that color features change with different video formats. Another drawback of color features is that they are not applicable to black and white videos. Color-space-based fingerprint

or color signature encodes the absolute color of the frames while discarding spatial information. If some movies contain the number of shots in different parts with the same color scheme then it is not possible to differentiate them using color-space-based fingerprint. Reason behind this is due to lack of spatial information.

Temporal Finger Prints

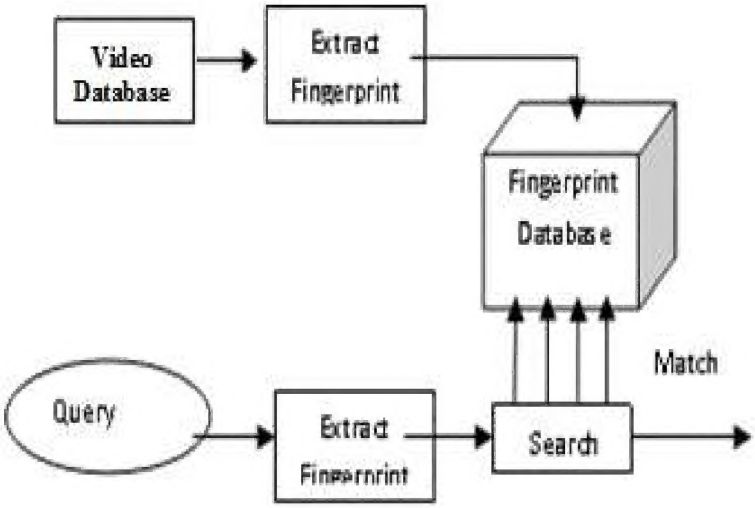

Figure 1: Structure of a Fingerprinting System

To overcome drawback of color-space-based finger prints new video fingerprint extraction algorithm is developed that can be applied to the luminance (the gray level) value of the frames. This technique needs a shot-boundary-detection algorithm, and it can be efficient for finding a full movie, but may not work well for short episodes with a few boundaries. The common way for shot detection is to evaluate difference value between consecutive frames represented by a given feature. First, a video sequence is segmented into shots. Then, the duration of each shot is taken as a temporal signature, and the sequence of concatenated shot durations form the fingerprint of the video. The key-frame based schemes are not robust to compression and resolution change, while the frame-by-frame based schemes are not robust to frame rate change, as well as that this type of signatures will be very large and has numerous redundant information. Therefore temporal ordinal based signatures are used.

Spatial Finger Prints

Spatial fingerprints are features derived from each frame or from a key frame. Spatial fingerprints can be further subdivided into global and local fingerprints.

Global Finger Prints

Global fingerprints focus on the global properties of a frame or a subsection of it like image histograms.

Local Finger Prints

It usually represents local information around some interest points within a frame like edges, corners etc. The scale invariant feature transform (SIFT) is an algorithm which is used to detect and describe local features in images. Key stages of SIFT includes Scale invariant feature detection, Feature matching and indexing, Cluster identification by the Hough transform voting, Model verification by linear least squares, Outlier detection

Spatio-Temporal Fingerprints

The new fingerprinting approach is designed to call as spatio-temporal fingerprints. Spatio-temporal finger prints that contain both spatial and temporal information. Spatial information describes the physical location of objects and metric relationship between objects. Temporal information is related to time. Some spatio-temporal algorithms consider a video as a three-dimensional (3-D) matrix and extract 3-D transform-based features. Discrete Cosine Transform (DCT) based hash algorithm and Randomized Basis Set Transform (RBT) algorithm.

PROPOSED METHOD

Figure 1 shows the overall structure of fingerprinting system. When an identifier or the signature is extracted from the content without changing the content, it is fingerprinting. Video fingerprinting has been used to refer to the technology encompassing algorithms, systems, and workflows that use video fingerprint for video identification

There are some disadvantages of existing fingerprint extraction systems .They are as follows

- 3-D transform to a video is a computationally demanding process.
- The computational bottleneck is the search time in the matching process rather than the fingerprint extraction time.
- Overlapping reduces the sensitivity of the fingerprints to the synchronization problem.
- The problem with the binarization scheme limits the number of coefficients and thus the fingerprint length that can be used.
- 3D-DCT is resistant to different types of distortions that can happen to video signals.

These drawbacks are overcome in proposed Temporally Informative Representative Images-Discrete Cosine Transform (TIRI-DCT) system. As a Temporally Informative Representative Image (TIRI) contains spatial and temporal information on a short segment of a video sequence, the spatial feature extracted from a TIRI would also contain temporal information. Based on TIRIs; an efficient fingerprinting algorithm is proposed to call as Temporally Informative Representative Images-Discrete Cosine Transform (TIRI-DCT).In TIRI – DCT, first step is generation of temporally informative representative images (TIRI. So proposed TIRI-DCT method along with fast search algorithm outperforms than 3D-DCT since it is more robust, discriminant, and fast.

Advantages of proposed system are as follows.

- A Spatio-temporal fingerprint is adopted because of their comprehensiveness.
- The TIRI-DCT method introduces a fingerprinting system that is robust, discriminant, and fast.
- The TIRI-DCT outperforms the well-established (3D-DCT) algorithm and maintains a good performance for different attacks like noise, time shift, spatial shift, brightness/contrast, rotation, frame loss that normally occurs on video signals.
- TIRI –DCT algorithm is applied to color videos as well as black and white videos.

DESIGN STEPS

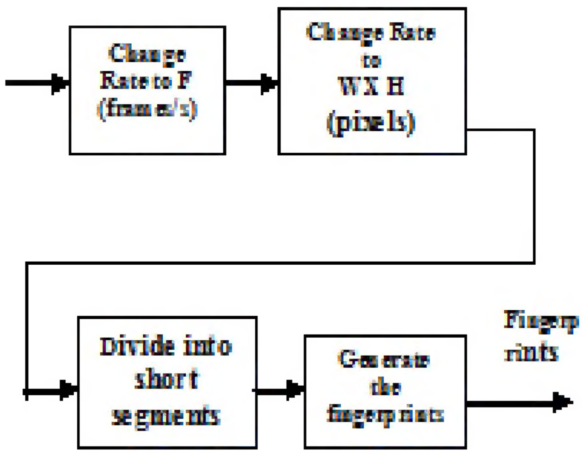

Figure 2: Preprocessing Steps

As shown in figure 2, each video is down-sampled both in time and space. Prior to down-sampling, a Gaussian smoothing filter is applied in both domains to prevent aliasing. This down-sampling process provides the fingerprinting algorithm with inputs of fixed size (W × H) pixels and fixed rate (F frames/second. After preprocessing, the video frames are divided into overlapping segments of fixed-length, each containing J frames. The fingerprinting algorithms are applied to these segments. Overlapping reduces the sensitivity of the fingerprints to the "synchronization problem" which is called as "time shift". Features are derived by applying a 2D-DCT on overlapping blocks of size from each TIRI. As shown in figure 2 the first horizontal and the first vertical Discrete Cosine Transform (DCT) coefficients (features) are extracted from each block. The value of the features from all the blocks is concatenated to form the feature vector. Each feature is then compared to a threshold (which is the median value of the feature vector) and a binary fingerprint is generated.

TIRI-DCT Algorithm

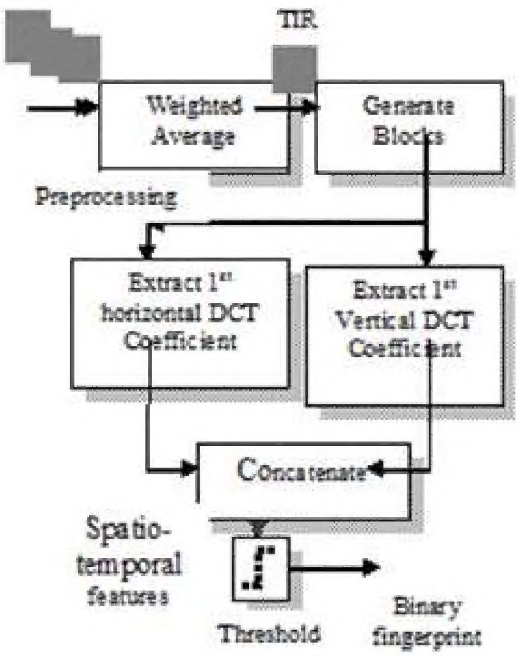

Figure 3: Schematic of the TIRI-DCT Algorithm

Figure 3 shows the block diagram of proposed approach which is based on temporally informative representative images (TIRIs). Preprocessing is often used to improve visual quality and coding efficiency of video compression systems. In signal processing, down sampling or sub- sampling is the process of reducing sampling rate of the signal. Since down sampling reduces the sampling rate, we must be careful to make sure the Shannon-Nyquist sampling theorem criterion is maintained. If the sampling theorem is not satisfied then resulting digital signal will have aliasing .Anti-aliasing means removing signal components that have a higher frequency than is able to be properly resolved by recording (or sampling) device. This removal is done before (re) sampling at a lower resolution. When sampling is performed without removing 5this part of the signal, it causes undesirable artifacts such as black and white noise. If the original signal had been bandwidth limited, and then first sampled at a rate higher than Nyquist minimum, then the down sampled signal may already be Nyquist compliant, so the down sampling can be done directly without any additional filtering. Down sampling only changes the sample rate not the bandwidth of the signal. The only reason to filter the bandwidth is to avoid the case where the new sample rate would become lower than the Nyquist requirement and then cause the aliasing by being below Nyquist minimum.

EXPECTED RESULTS

Test Videos

Temporally informative representative images- discrete cosine transform (TIRI-DCT) is spatio-temporal fingerprint extraction algorithm. So it can be applied to both black and white videos as well as color videos. Video extension is .yuv with QCIF format having resolution of$176 \times 144 (W \times H)$.

Following black and white videos are used for testing.

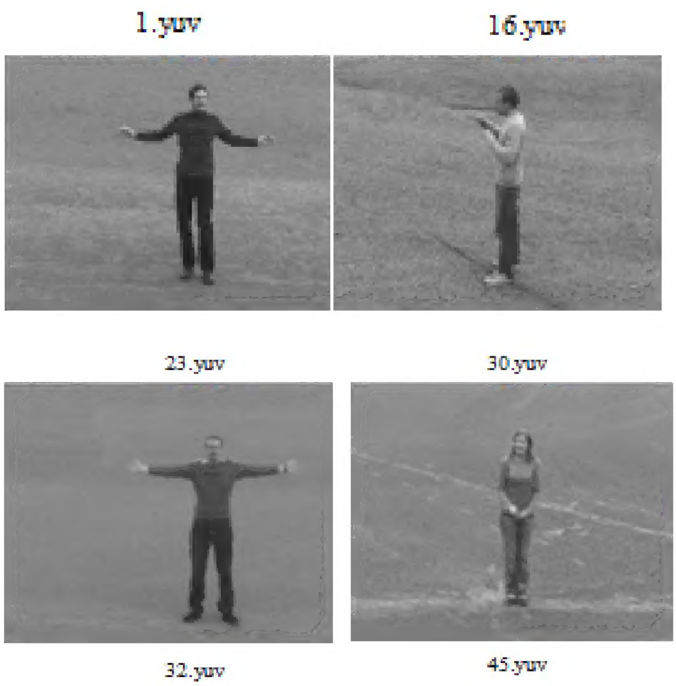

Figure 4

Results for Different Weighting Factors

Results for different weighting factors like constant, linear and exponential are as given below.

Constant, linear and exponential weighting factors for video 1.yuv:

(a) 189 th Frame of Input Video 1. Yuv

(b) Constant

(c) Linear

(d) Exponential

Figure 5: Frame (a) 189 th Frame of Input Video 1. yuv and Resulting TIRIs with Different Weighting Factors (b) Constant (c) Linear (d) Exponential

RESULTS FOR F-SCORE

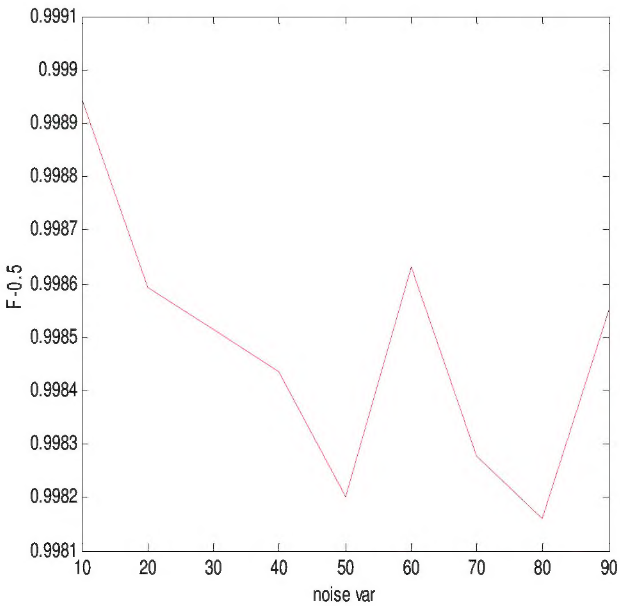

Figure 6(a): F-Score vs Noise

From figure 6(a) it observed that even if the noise is increased from 10 to 70, F-Score value is not randomly decreased. For noise range from 10 to 70 F-Score value is closer to 1 which indicates better performance of the TIRI-DCT system.

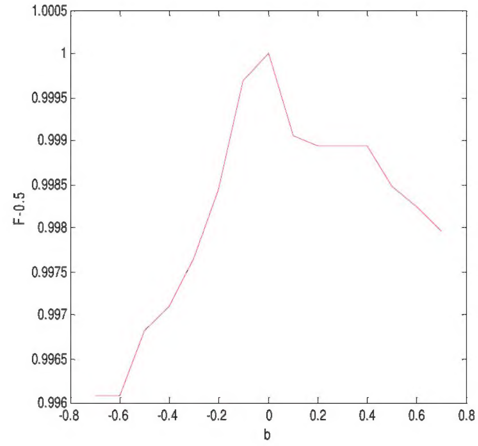

Figure 6(b): F-Score vs Brightness

Figure 6(b) shows that when brightness is increased from -0.6 to 0 then F-Score increases up to 1. Once F-Score reaches to 1 and brightness of video is again increased then F-Score performance is not degraded.

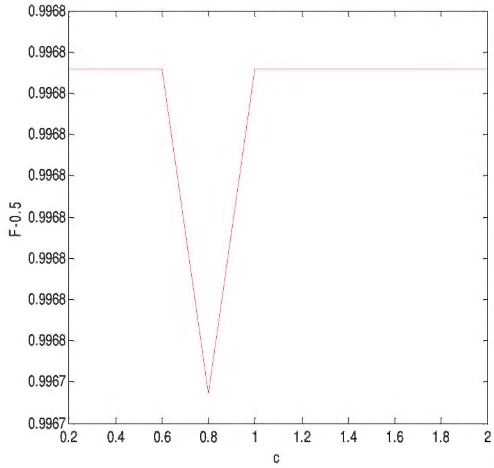

Figure 6(c): F-Score vs Contrast

Figure 6(c) Indicates that for Contrast Range from 0.2 to 2 F-Score is nearly remains constant and F-Score value is closer to 1

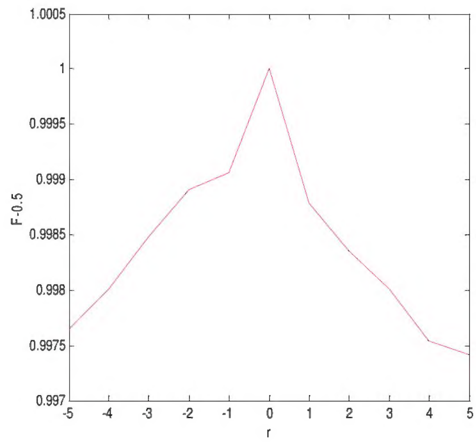

Figure 6(d): F-Score vs Rotation

From Figure 6(d) it observed that if video frames are not rotated then TIRI-DCT maintains very good performance in terms of the F - Score. But if frames are rotated in negative or positive degree then performance degrade slightly.

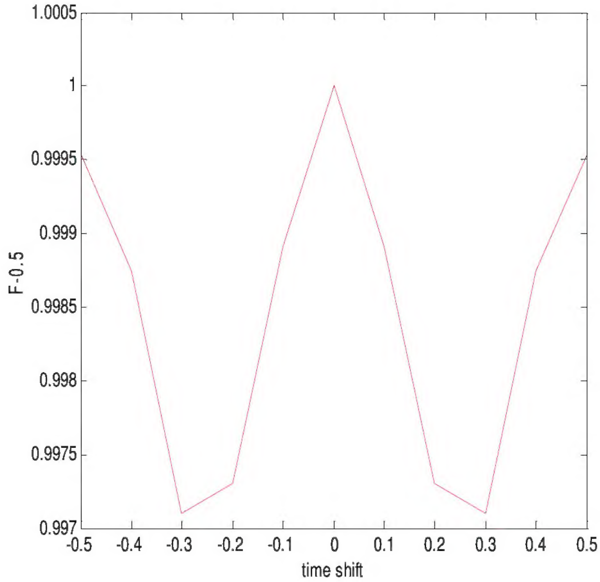

Figure 6(e): F-Score vs Time Shift

Figure 6(e) shows that if the video is shifted by some second from 0 to 0.5 seconds or from 0 to –0.5 seconds then F-Score value is slowly decreased but still it is closer to 1, representing good performance of TIRI-DCT. If video is not shifted in time, means the beginning of query video is exactly aligned with beginning of reference video then F-Score value reaches to 1.

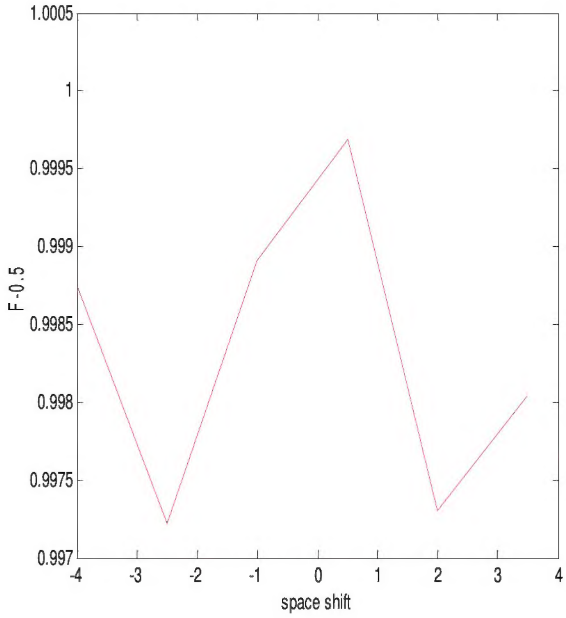

Figure 6(f): F-Score vs Space Shift

Figure 6(f) concludes that if a video frame is shifted by -4 (%) to 4 (%) right and-4 (%) to 4 (%) down then F-Score maintain good performance.

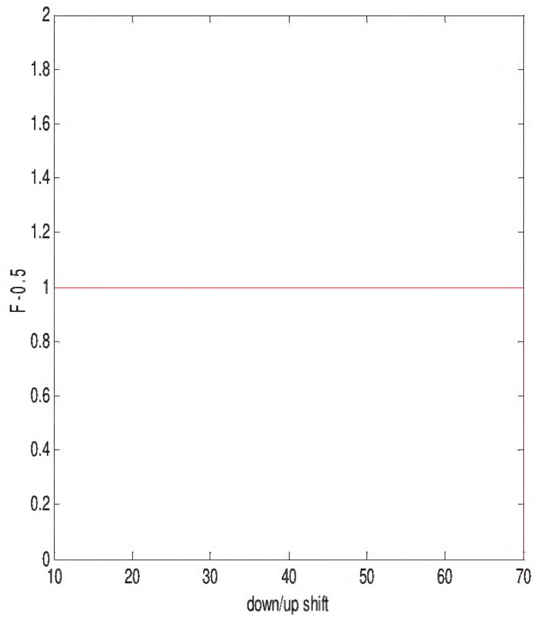

Figure 6(g): F-Score vs Frame Loss

Figure 6: F-Score of TIRI -DCT for Different Attack Parameters (a) Noise (b) Brightness (c) Contrast (d) Rotation (e) Time Shift (f) Spatial Shift (g) Frame Loss

For all test videos TIRI–DCT maintains good performance over attacks like noise, brightness, contrast, rotation because the F-Score is closer to 1, which indicates the perfect classification system.

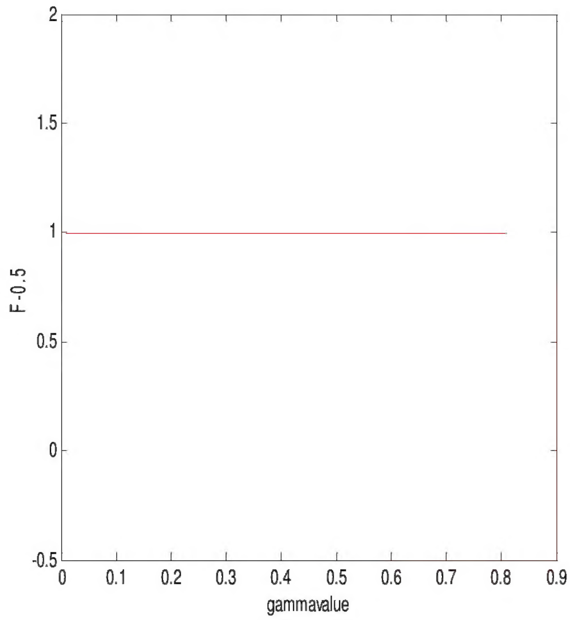

Figure 7

CONCLUSIONS

This paper proposes a fingerprinting system for video copy detection. It can be used for copyright management and indexing applications. this paper to discuss robustness, discrimination, security, and fast search of fingerprints simultaneously. The system consists of a fingerprint extraction algorithm followed by an approximate search method. The proposed fingerprinting algorithm (TIRI-DCT) extracts robust, discriminant, and compact fingerprints from videos in

a fast and reliable fashion. These finger prints are extracted from TIRIs containing both spatial and temporal information about a video segment. We demonstrate that TIRI-DCT generally outperforms the well-established (3D-DCT) algorithm and maintains a good performance for different attacks on video signals, including noise addition, changes in brightness/contrast, rotation, spatial/temporal shift, and frame loss.

FUTURE WORK

As part of our futurework, we will conduct a detailed analytical study of the security of fingerprinting algorithms including the one proposed in this paper. As another part of our future work, we will carry an extensive comparison study to compare our fingerprinting algorithms to other state-of-the-art algorithms. We will also evaluate our proposed fast search methods when applied to other fingerprinting methods. We also plan to study the performance of the system in the presence of some other attacks, such as cropping, and logo insertion.

REFERENCES

1. IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 6, NO. 1,MARCH 2011 213A Robust and Fast Video Copy Detection System Using Content-Based Fingerprinting by-Mani MalekEsmaeili, Mehrdad Fatourechi, and RababKreidieh Ward*, Fellow, IEEE*

2. S. Lee and C. Yoo, "Robust video fingerprinting for content-based video identification," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 7, pp. 983–988, Jul. 2008

3. C. De Roover, C. De Vleeschouwer, F. Lefebvre, and B. Macq, "Robust video hashing based on radial projections of key frames," *IEEE Trans. Signal Process,* vol. 53, no. 10, pp. 4020 4037, Oct. 2005.

4. S. Lee and C. Yoo, "Robust video fingerprinting for content-based video identification," *IEEE Trans. Circuits Syst. Video Technol.*, vol.18, no. 7, pp. 983–988, Jul. 2008.

5. X. Su, T. Huang, and W. Gao, "Robust video fingerprinting based on visual attention regions," in *Proc. ICASSP*, Washington, DC, 2009, pp.1525–1528, IEEE Computer Society.

6. Joly, O. Buisson, and C. Frelicot, "Content-based copy retrieval using distortion-based probabilistic similarity search," *IEEE Trans. Multimedia*, vol. 9, no. 2, pp. 293–306, Feb. 2007.

7. M. Malekesmaeili and R. K. Ward, "Robust video hashing based on temporally informative representative images," in *Proc. IEEE Int. Conf. Consumer Electronics*, Jan. 2010,